

What happens if you retest autobiographical memory 10 years on?

CHRISTOPHER D. B. BURT and SIMON KEMP
University of Canterbury, Christchurch, New Zealand

and

MARTIN CONWAY
University of Bristol, Bristol, England

Burt (1992a, 1992b) reported data on the autobiographical memory of diarists for events that had occurred on average 3.3 years earlier. This paper reports data on 11 of the diarists, who were recontacted after a further 10 years and who agreed to a retest of their memory. Estimates of event date and event duration from the two recall attempts were compared. As predicted, duration estimation was extremely stable and showed no detrimental effects of the additional 10 years of retention interval. Estimation of event date was predicted to show an increase in forward telescoping due to the increased remoteness of the event sample, but, contrary to this prediction, backward telescoping dominated dating errors. A combination of the establishment of a recent boundary and Kemp's (1999) associative model of dating is proposed as an explanation for these results. It is argued that the nature of dating errors may depend on the time of the event's occurrence in the life span and the age of the individual dating the events.

Over the past 2 decades there have been several intensive examinations of autobiographical memory (e.g., Burt, 1992a, 1992b; Linton, 1978; Wagenaar, 1986). One of the dominant features of this work is its reliance on cross-sectional data sets. Although memory for events with vastly different retention intervals has been examined, recall has typically been at one point in time. Despite this, there have been many conclusions drawn about the causal relations between elapsed time and autobiographical memory processes. In particular, research on the retrieval of the time of an event's occurrence has devoted considerable effort to identifying how error patterns might vary with event age and what processes might explain such errors. The research reported in this paper examined temporal memory error patterns and tested theories of temporal memory using a within-subjects repeated measures design that provides a unique view of the influence of elapsed time on temporal memory. A group of diarists whose memory for events was first examined in 1987–1988 (referred to as Test 1) was contacted and retested in 1997–1998 (referred to as Test 2). Their ability in the two tests to estimate the duration of their experiences and date when events occurred is compared.

Two temporal aspects of memory were central to this study: the estimation of event duration and the retrieval of

date information. We consider duration estimation first. There is a long history of research on both memory for and the perception of duration (see Fraisse, 1984, for a review). Much of this work has examined duration estimation under controlled laboratory conditions using stimuli such as clicks (e.g., McConchie & Rutschmann, 1971), electrical current (e.g., Ekman, Frankenhaüser, Levander, & Mellis, 1966), a moving spot (e.g., Rachlin, 1966), music (e.g., Kowal, 1987), and performance of a mathematical task (e.g., Burnside, 1971).

In contrast to this laboratory research there is a limited number of studies that have required individuals to estimate the duration of everyday events. Schneider, Griffith, Sumi, and Burcart (1978) asked crime victims to estimate the time it took for the police to arrive at the crime scene and found estimates were generally longer than those recorded. Baddeley, Lewis, and Nimmo-Smith (1978) asked members of an experimental psychology research panel to recall the duration of previous visits to the laboratory and reported estimates that were not significantly different from the actual durations. Douglas and Blomfield (1956) used hospital records to assess the accuracy of mothers' recall of the duration of their child's hospitalization. A correlation of .94 was found between actual and estimated duration. Mednick and Shaffer (1963) compared pediatricians' records with mothers' estimates of their breast-feeding duration and found a tendency toward overestimation. Pyles, Stolz, and MacFarlane (1935) used archival records obtained from hospitals, physicians, and public health nurses to assess the accuracy of mothers' estimates of the duration of labor, obtained approximately 21 months after birth. A correlation of .61 was

This research was supported by a grant from the Marsden Fund (MS1012). We thank Rich Marsh, Charles Thompson, and anonymous reviewers for commenting on earlier versions of this manuscript. Correspondence should be addressed to C. D. B. Burt, Department of Psychology, University of Canterbury, Private Bag 4800, Christchurch, NZ (e-mail: cburt@psyc.canterbury.ac.nz).

obtained between estimates of duration of labor and recorded duration.

The initial (Test 1) examination of the duration estimation abilities of the participants in this study (e.g., Burt, 1992a) used autobiographical events and compared estimates with records showing the actual duration of the events. That work, and a number of subsequent studies (e.g., Burt 1993, 1999; Burt & Kemp, 1991, 1994; Burt & Popple, 1996) concluded that duration estimation for everyday events involved considerable reconstructive processing. The reconstruction of event duration is suggested to begin with the classification of the target event as belonging to a particular category. This classification provides access to general knowledge that the individual has regarding the typical characteristics of that type of event. For events that have a typical duration (e.g., a holiday normally lasts for a specific period of time), this temporal attribute would probably be available. It was shown in Burt and Kemp (1991) that increased experience with an event category reduced the between-subjects variation in duration estimates, suggesting that with experience individuals tend to agree more as to the typical duration of particular types of events. The accuracy of reconstructed duration estimates that are based on retrieval of typical event duration was shown in Burt (1993) to depend on the typicality of actual event duration and the individual's ability to use his/her memory of the event to adjust his/her reconstructed estimate for atypical event characteristics.

Such an explanation of the role of event memory in duration estimation is in contrast to a number of other models of duration estimation, the most dominant of which are Ornstein's (1969) storage size hypothesis and Block and Reed's (1978) contextual change hypothesis. The storage size hypothesis suggests that the amount of information that is stored in memory about an interval is positively related to how long that interval is estimated to be. In contrast, Block and Reed argued that estimated duration is related to the number of stimulus changes occurring within an interval. A number of laboratory studies have supported these models, finding that an interval that contains complex, unfamiliar, less predictable, or more numerous components is generally estimated to be significantly longer than an interval of similar duration containing simpler (e.g., Block, 1978, Experiment 2; Schiffman & Bobko, 1974), more familiar (e.g., Avant & Lyman, 1975), more predictable (e.g., Ornstein, 1969), or fewer (e.g., Buffardi, 1971; Burnside, 1971; Poynter & Homa, 1983; Schiffman & Bobko, 1977; Thomas & Brown, 1974) components. The latter finding, relating to the number of components within the interval, is frequently referred to as the *filled-duration illusion*. Burt (1992a) found no influence of the amount of event content on estimated duration.

The evidence that a reconstructive model accounts rather well for the diarists' duration estimation abilities prompted the prediction that there would be very little change in their Test 2 estimates relative to the Test 1 es-

timates. Some variation might be expected if loss of memory for an event reduced the ability to adjust estimated duration for atypical event characteristics. However, this might be compensated for by the extra years of experience, which may have increased knowledge of typical event duration and thus made reconstructed estimates more accurate, particularly where actual event duration was typical.

We turn now to event date estimation. A number of studies have used single-word or phrase prompts to facilitate the recall of specific autobiographical events and then requested that each event be dated (e.g., Crovitz & Schiffman, 1974; Fitzgerald, 1981; Fitzgerald & Lawrence, 1984; Holding, Noonan, Pfau, & Holding, 1986; Pillemer, Goldsmith, Panter, & White, 1988; Pillemer, Rhinehart, & White, 1986). A difficulty with this method is the inability to compare estimated dates with the true dates, as these are not known. This difficulty essentially precludes the examination of dating error patterns. In contrast, the studies by Barclay and Wellman (1986); Bruce and Van Pelt (1989); Burt (1992b); Larsen and Thompson (1995); Linton (1975); Thompson (1982, 1985a, 1985b); Thompson, Showronski, and Betz (1993); Thompson, Showronski, and Lee (1988); Wagenaar (1986); and White (1982) achieved verification of actual event dates by using diary records. Other studies have used archival records to determine the participant's involvement in, and the precise actual date of, the events to be dated (Baddeley et al., 1978; Loftus & Fathi, 1985). Finally, there is a large body of dating research that has used public events (e.g., Brown, Rips, & Shevell, 1985; Brown, Shevell, & Rips, 1986; Ferguson & Martin, 1983; Friedman & Wilkins, 1985; Kemp, 1987, 1988; Kemp & Burt, 1998; Perlmutter, Metzger, Miller, & Nezworski, 1980). Although the events in this research are not autobiographical, the results show considerable consistency with the autobiographical data.

Two dating error patterns have consistently been identified. Dating error increases as retention interval increases (e.g., Barclay & Wellman, 1986; Burt, 1992b; Friedman & Wilkins, 1985; Linton, 1975; Perlmutter et al., 1980; Thompson, 1982, 1985a, 1985b; Thompson et al., 1988), and the sign of the event dating error—that is, whether the error is an under- or overestimation of an event's actual age—appears to be systematically related to retention interval. Brown et al. (1985) reported a product-moment correlation between signed error and actual date of $-.74$, and Baddeley et al. (1978) reported $-.33$. These correlations suggest a tendency toward underestimation of event age, or *forward telescoping*, as retention interval increases, a result that is generally found when the dating error of a number of events of different ages is averaged (Kemp, 1988). These results led to the prediction that the Test 2 dating signed errors might on average indicate that the events were dated as having occurred more recently (forward telescoping) when compared with the dates of occurrence estimated at Test 1.

This prediction is based on the reasoning that mean retention interval increased from 3.34 years at Test 1 to 13.27 years at Test 2, and therefore clearly the events were substantially older at Test 2 than at Test 1.

The latter prediction, that redating events after a relatively long test–retest interval will produce increased underestimation of event age, appears not to have been previously examined. The few test–retest dating studies that have been conducted required participants to recall a specific autobiographical experience cued by a prompt word (e.g., McCormack, 1979; Robinson, 1976, 1986). Thus there was no opportunity to examine dating error patterns. In McCormack's study, date was recalled by providing the calendar year, the duration since the event, or the age when the event occurred, while Robinson (1976, 1986) requested year and month as a minimum. Elapsed time between dating and redating also varied slightly between the studies; Robinson used an interval of 2 weeks in 1986 and 1 week in 1976, the latter time being similar to that used by McCormack (1979). All three studies found reasonably high product–moment correlation coefficients between the two sets of assigned dates (McCormack, 1979, $r = .98$; Robinson, 1976, $r = .94$; and Robinson, 1986, $r = .66$ and $.84$).

In addition to examining dating ability, researchers have focused on explaining the processes involved in dating. Friedman (1993) reviewed the dating literature and concluded, as other authors have (e.g., Kemp, 1999; Thompson, Skowronski, Larsen, & Betz, 1996), that date generation involves the application of a limited set of strategies. Both distance- and location-based dating strategies have been proposed (Friedman, 1996). Distance-based models suggest that date generation may be based on the strength of the memory trace for an event (see Bradburn, Rips, & Shevell, 1987; Brown et al., 1985; Kemp & Burt, 1998; Friedman & Kemp, 1998). Location-based processes allow an individual to make use of knowledge of behavior patterns, and the subset of events often referred to as *landmark events* (see Shum, 1998), for which the actual date can be recalled, to construct date information.

In summary, in the present study we contacted 11 diarists who, in 1987–1988, had dated and estimated durations for a sample of their experiences. These diarists were presented with the same events as in 1987–1988 and again asked to date them or estimate their duration. This within-subjects repeated measures analysis of autobiographical memory with a retest interval of approximately 10 years is unique and provides a powerful examination of a number of the processes that are assumed to characterize temporal memory. On the basis of a reconstructive model, the duration estimation ability of the participants at Test 2 was expected to be on average very similar to that shown at Test 1. In contrast, the dating of events was expected to show a number of changes: Absolute dating error was expected to increase, and the nature of the dating errors was expected to change with an increased tendency to forward telescoping. Dating strategies were also expected to

change. Access to landmark events was expected to decline, with an associated increase in the use of other strategies.

METHOD

Participants

Fourteen diarists participated in Test 1 (see Burt, 1992a and 1992b, for a full description of the Test 1 method). Eleven of these were re-located for the present study (Test 2). Three of the participants were male, all 31 years of age at Test 2, while the mean age of the female participants was 30.7-years (range = 29–34 years) at Test 2.

Diary Records

Prior to Test 1 the diarists had submitted their diary records to the first author for examination. Each diary was a reasonably complete account of at least a year's events, with a record entered for each day. The 11 diarists located at Test 2 originally submitted 41 years of diary material covering the period from 1979 to 1987. At the time of Test 1 (1987–1988), events were extracted from the diary records and developed into an interview schedule.

The participant's diaries had been returned after Test 1. At the time of recontacting the diarists, each was asked whether he/she still had those diaries: all but 1 participant had retained them. The participants were also asked when they had last examined them. Seven indicated the last examination was when they were returned in 1988. Of the other 3 participants, 1 did not respond to the question, and for the other 2 it was 4 and 7 years ago, respectively.

Two types of event (labeled *specific* and *duration events*) were extracted from the diaries and were examined at both Test 1 and Test 2. Specific events were defined as having occurred on a particular day or date. To be considered for selection from the diary material, four attributes had to be determinable: the *what*, *where*, *who*, and *when* of the event (see Burt, 1992b, for further discussion of each attribute).

The defining characteristic of a duration event was that it had lasted for more than 1 day. Two types of duration events were obtained: "filled" and "empty" duration. Empty duration events consisted of two associated events separated by some period of time (e.g., an individual sends away for a rare stamp and later receives the stamp). The interval between the two events is *assumed* to be relatively uncontaminated by associated events. Filled duration events consist of a sequence of specific events that were associated because they occurred within a definable interval of time, such as a holiday. Actual duration was obtained from the diary records with both the day on which the event began and that on which it ended counted as a whole day.

Questionnaire Construction

Because the participants now lived in a number of different countries, Test 2 was performed using questionnaires in contrast to the interview method adopted for Test 1. All the events for which data had been collected at Test 1 were included in each participant's Test 2 questionnaires, including events that had not been recalled or recognized at Test 1. Events (both specific and duration) were assigned randomly to a questionnaire page: a full event description headed the page (e.g., *You and Justin made a fire and bivouac, and toasted marshmallows at the Rakaia river*). For specific events this was followed by a 9-point memory rating scale and a place to enter the event's date (day, month, and year) and the strategy adopted to generate each date component.

The 9-point specific event memory rating scale was tagged at each point and gave an opportunity to indicate different types of memory failure or degree of recall (1 = *it never happened*, 2 = *the combination of cues does not prompt any recollection*, 3 = *it could be one of many similar events*, 4 = *I can just barely remember it*, 5 = *I remember it but not so well*, 6 = *I remember it fairly well*, 7 = *I*

remember it very well, 8 = I remember it almost perfectly, 9 = I remember it perfectly).

The strategy used to generate each date component (day, month, and year) was indicated by selecting from five choices (combinations could be used): 1 = remembered date component, 2 = guessed date component, 3 = reconstructed date component using events that I know occurred *before* the target event, 4 = reconstructed date component using events that I know occurred *after* the target event, 5 = reconstructed date component using specific information provided by the description, e.g., the event obviously occurred in summer, winter, May holidays, etc.).

An example of a duration event description was as follows: *You rang Richard M. and asked him if he would go to the school formal with you. He said he needed more time to think. Sometime later he rang back and said he had an exam and couldn't go.* These descriptions were followed by a 7-point memory scale (1 = I cannot remember this event, 2 = I can just barely remember it, 3 = I remember it but not so well, 4 = I remember it fairly well, 5 = I remember it very well, 6 = I remember it almost perfectly, 7 = I remember it perfectly). Three other 7-point rating scales were used (confidence in the accuracy of your duration estimate—1 = no confidence, 7 = absolute confidence; distinctiveness of the event, 1 = not at all distinct, 7 = very distinct; and importance of the event, 1 = of no importance, 7 = very important).

False Event Construction

A number of false events were embedded within each participant's questionnaire/s. The participant's true events formed the base for the false events. In order to transform a true event into a false event, one or more of its attributes (*what*, *where*, and *who*) was changed. The component/s that were used as replacements were taken from other events within the participant's sample of events. Thus, when we created false event descriptions, these featured people that the participant had interacted with, locations that they had been to, and activities that they had engaged in. Seven event transformations were created by changing activity only, location only, participants only, activity and location, activity and participant, location and participant, and activity, location, and participant. The last type of transformation created a completely false, yet plausible, event. An attempt was made to equate the number of the different types of event transformations. Each false event description was checked for plausibility. For example, the false event *You and dad went to a fancy dress party at Donna's place* would have been rejected. The false events are briefly mentioned in the Results and General Discussion sections; for a more in-depth discussion of the participants' responses to the false events and related experiments see Burt, Kemp, and Conway (2000).

Procedure

Each questionnaire was limited to a maximum of 70 pages, which was the maximum we felt the participant could reasonably be expected to complete in one sitting. This was posted to the participant. The number of questionnaires completed by the participants ranged from one to five. Each questionnaire was sealed in an envelope and enclosed with an instruction sheet in another envelope. These instructions asked participants to ensure they had sufficient time available to work on the questionnaire before they opened it. It was noted that they should not look over the questionnaire before working on it, should start at the beginning and complete the questions (events) in the order presented, should not go back and change an answer (there was no evidence to suggest this had occurred), should use *only* their memory to answer the questions, and should take breaks when they needed to.

Finally, the date on which they completed each questionnaire page (event) and how long they required to read the event description and complete the responses on the page were entered at the top of each page. The participant's next questionnaire was posted approximately 1 month after we received a completed questionnaire.

RESULTS

The Test 1 data file was matched to the Test 2 data file. Events were excluded if there were excessive missing data. For this reason, and because of the loss of several participants, there are differences between the overall Test 1 results reported here and those reported in Burt (1992a, 1992b). Test 1 and Test 2 matched data were examined for 576 specific events and 144 duration events. The analyses below are divided into a number of sections: memory for events, dating of events, dating strategies, and recall of event duration. A number of the analyses used events as the random variable and thus assumed a degree of independence between the memories (see Wright, 1998, for a discussion of this issue). On the basis of our previous examinations of autobiographical memory data sets for individual differences (see Burt, Watt, Mitchell, & Conway, 1998), we did not expect individual participant differences to bias the results in this study. However, culling of the data set was performed in order to demonstrate that some of the key findings were not an artifact of a specific participant's data.

Memory for Events

Memory for the duration events was generally rated higher than that for the specific events, with 81.6% of the memory ratings of the former indicating that the participant felt he/she remembered the event, compared with only 55.0% of the specific event ratings. Of the remaining specific events ($n = 259$), 41.3% were rated as too similar to other events to be remembered, 45.1% could not be recalled because insufficient information was provided in the event description, and 13.6% were rated as never having happened. Of the 514 false events, 21.5% were rated as being remembered.

Of interest was the comparison of Test 1 and Test 2 event memory. In contrast to Test 2, where event memory was expressed as a rating, event memory was objectively assessed using an event cuing procedure at Test 1 (see Burt, 1992b, for details), and events were assigned to one of six categories: no memory for event, event recognized after all cues were given, and four levels representing increased ability to recall the event determined by a cue efficiency rating. As a means of comparing event memory across Test 1 and Test 2, mean Test 2 memory ratings were calculated for each of the six categories of event memory defined by Test 1 performance: respective means of 2.85, 3.39, 4.36, 4.79, 4.13, and 4.36 were obtained for the events in each of the six memory categories. Inspection of these means indicates that they increase with the memory measure taken at Test 1 (e.g., the events that were not remembered at Test 1 received, on average, the lowest Test 2 memory rating). Comparison of Test 2 memory ratings across the six Test 1 memory categories indicated a significant difference [$F(5,570) = 18.827$, $p > .01$]. In other words, there was at least some consistency in the participant's memory for the events across the two tests. However, there was some switching between forgetting and remembering from Test 1 to Test 2, with

4.9% of the events that were not remembered at Test 1 rated as remembered at Test 2.

In order to correlate the Test 1 and Test 2 event memory data, the Test 1 specific memory categories were recoded so as to assign a Test 1 memory value to each event (events that were not remembered at Test 1 = 1, events recognized after all cues were given = 2, and the four levels representing increased ability to recall the event determined by a cue efficiency rating were assigned 3, 4, 5, and 6, respectively). A significant association was obtained when the Test 1 (recoded) and Test 2 memory ratings were correlated ($r = .34, p < .01$). Test 1 memory for duration events was rated on the same 7-point scale as used at Test 2. Overall, the memory ratings given for the duration events at Test 1 and Test 2 were found to be significantly correlated ($r = .35, p < .01$). Comparison of the memory ratings given at Test 1 ($M = 3.9$) and Test 2 ($M = 3.1$) indicated a significant decrease in event memory [$F(1,143) = 30.805, p < .001$]. The moderate level of these correlations indicates a marked degree of change in the memorability of both the specific and duration events over the period between the two tests.

Dating of Events

Participants were requested to provide a date for each event irrespective of its memory rating. Examination of the relationship between Test 2 absolute dating error (actual date - Test 2 assigned date, and ignore the sign), and the memory ratings produced a small but significant correlation ($r = -.10, p < .05$). Furthermore, comparison of the absolute dating error for each of the nine memory rating categories indicated significant differences [$F(8,564) = 2.544, p < .01$]. Inspection of the mean absolute dating error for each memory rating category indicated that the least accurate dating was for the events rated as *never having happened*. Exclusion of these events removed the significant difference when the absolute dating errors of the remaining memory rating categories were compared and reduced the correlation between the memory ratings and absolute dating error ($r = -.06, n.s.$). Thus it seemed reasonable to consider the events

rated as *never having happened* to be “outliers,” and they were excluded from the remaining analyses.

Table 1 shows the correlations between actual date and estimated date at Test 1 and Test 2 for each participant. Also shown are the correlations between the Test 1 and Test 2 estimated dates for each participant. Inspection of Table 1 suggests that, as expected, the dating ability deteriorated from Test 1 to Test 2. Generally the correlations between Test 2 estimated date and actual date were lower than those between the Test 1 estimated date and actual date. However, column 5 of Table 1 shows that a number of the participants had significant positive correlations between their Test 1 and Test 2 dating. Table 1 also indicates that 1 participant contributed 179 events to the sample and thus could potentially have had a marked influence on the overall results. However, with data from this participant excluded ($n = 362$), the overall correlations remained essentially unchanged (actual date and Test 1 estimated date $r = .92, p < .01$; actual date and Test 2 estimated date $r = .83, p < .01$; and Test 1 and Test 2 estimated date $r = .81, p < .01$).

Absolute dating error was calculated for both Test 1 and Test 2 estimated dates (as described above). Comparison of the Test 1 and Test 2 absolute dating errors indicated a significant difference [$F(1,540) = 58.521, p < .001$], with the Test 2 mean absolute error (281.6 days) greater than the Test 1 mean (166.1 days). To ensure that this large difference was not due to a few extremely inaccurate dates at Test 2, the comparison was repeated with the data set culled of all absolute errors greater than 2 SD from the Test 2 absolute error mean. This reduced the data set by 29 events but did not change the significant difference between the Test 1 and Test 2 absolute errors [$F(1,511) = 30.457, p < .001$; new Test 1 mean = 162 days and new Test 2 mean = 231.4 days].

An error difference score was calculated by subtracting the Test 2 absolute error from Test 1 absolute error (for all 541 events). Table 2 shows the mean error difference score for each participant. A negative error difference score indicates that on average the participant was less accurate at dating events at Test 2, and as predicted this was the case overall and for 9 of the 11 participants. Interestingly, 2 participants showed an average improvement in their dating accuracy.

Mean signed dating errors (actual date - test x assigned date) were calculated for each participant at Test 1 and Test 2 and are also shown in Table 2. A positive signed error indicates that the event was dated too remotely (the date was backward telescoped). As indicated in the second column of Table 2, the overall Test 1 signed error mean was consistent with that obtained in many other dating studies and showed a clear trend toward events being forward telescoped. In contrast, the Test 2 overall signed error mean (and eight of the individual means) showed a trend toward backward telescoping. Comparison of the Test 1 and Test 2 signed errors indicated a significant difference [$F(1,540) = 23.924, p < .001$]. This analysis was repeated with the data set culled of the 29 events, with Test 2 absolute errors greater than 2 SD

Table 1
Mean Retention Interval and Dating Correlations for Each Participant

| N | Mean Test 2 Retention Interval (Years) | Actual Date and Test 1 | Actual Date and Test 2 | Tests 1 and 2 |
|-------------|--|------------------------|------------------------|---------------|
| 48 | 11.9 | .71* | .25 | .17 |
| 16 | 11.0 | .91* | .52* | .59* |
| 14 | 12.8 | -.17 | .51 | .45 |
| 22 | 13.1 | .72* | .74* | .47* |
| 179 | 13.9 | .92* | .90* | .85* |
| 41 | 10.8 | .87* | .14 | .17 |
| 96 | 15.2 | .75* | .52* | .46* |
| 62 | 12.4 | .71* | .35* | .41* |
| 20 | 11.7 | .69* | .45* | .25 |
| 18 | 11.1 | .74* | .37 | .45 |
| 25 | 13.0 | .90* | .71* | .62* |
| Overall 541 | 13.3 | .93* | .84* | .82* |

* $p < .01$.

Table 2
Mean Absolute Error Differences and
Mean Signed Dating Error (Days)

| Tests 1 and 2 Mean Absolute Error Difference | Mean Signed Error | |
|--|-------------------|--------|
| | Test 1 | Test 2 |
| -154.3 | -67.1 | -100.3 |
| -48.1 | -3.6 | -11.3 |
| 42.3 | 25.1 | 32.8 |
| 14.9 | -261.7 | 186.8 |
| -62.7 | -2.9 | 23.3 |
| -53.6 | 11.3 | 5.4 |
| -174.4 | 2.5 | 172.6 |
| -182.5 | -16.7 | 140.1 |
| -527.4 | -33.9 | 463.5 |
| -71.9 | -38.3 | -56.2 |
| -71.9 | -77.7 | 53.6 |
| Overall -115.3 | -23.7 | 71.7 |

from the mean [$F(1,511) = 11.934, p < .001$; new Test 1 signed error mean = -27.0 days and new Test 2 signed error mean = $+32.6$ days]. The data set was *further* culled by including only those events that the participants rated as remembered at Test 2 and the analysis repeated [$F(1,306) = 9.783, p < .01$; new Test 1 signed error mean = -15.0 days and new Test 2 signed error mean = $+50.9$ days]. Finally, the data set was *further* culled through the exclusion of the participant with the largest single contribution to the data and the analysis repeated [$F(1,217) = 14.926, p < .001$; new Test 1 signed error mean = -18.5 days and new Test 2 signed error mean = $+80.1$ days]. Thus it would appear that there was an overall tendency for dating errors to be dominated by backward telescoping at Test 2.

As noted, distance-based models of dating suggest that memory strength can be used as a basis for date generation. Signed dating error was thus correlated with the recorded Test 1 event memory categories ($r = .02$) and with the Test 2 memory ratings ($r = .06$). The obtained correlations suggest that there was little influence of memory strength on the participants' date estimates.

The actual dates of the events occurred within "boundaries" defined by the beginning and end of the participant's diary keeping. These recent and remote boundaries were unique to a participant (although several participants had supplied diary material covering the same years). The Test 1 and Test 2 estimated dates ($n = 541$) were examined against the participant's recent and remote boundaries, and the proportion of events estimated as occurring outside each boundary was calculated. No events were dated more recently than the appropriate recent boundary at Test 1, and only seven (1.2%) events were dated more recently than the recent boundary at Test 2. Fifteen events (2.7%) had been dated more remote than the appropriate remote boundary at Test 1, and this increased to 39 events (7.2%) at Test 2.

Dating Strategies

No dating strategy information was provided for six events. Cross-tabulation of the dating strategies used for

the generation of each date component at Test 1 and Test 2 indicated that a number of event components were dated using the same strategy at Test 1 and Test 2. Overall, 72% of the day components, 45.4% of the month components, and 33.8% of the year components were generated using the same strategy at Test 1 and Test 2. In terms of the specific strategy used, 36.5% of the date components that were remembered at Test 1 were also remembered at Test 2, 83.3% were guessed on both tests, 39.3% were dated using landmark events on both tests, and 17.8% were dated using cue information on both tests.

It was predicted that the use of landmark events might show a particular decline at Test 2, as the passing of a further 10 years might have made these events harder to recall. At Test 1, 54 event components were dated using landmark events occurring before the event to be dated, and 252 were dated using landmark events occurring after the event to be dated. At Test 2 there were 71 instances where landmarks occurring before the event to be dated were used, and 133 where the landmark event occurred after the event to be dated. Thus, the participants may have been using different landmark events at Test 2, and drawing these from further back in their past.

Recall of Event Duration

Table 3 shows the Test 1 and Test 2 duration estimation means for each participant. Inspection of the individual means in Table 3 indicates that average Test 2 estimates were very similar to average Test 1 estimates (and to the mean actual duration). Indeed, the similarity between the overall mean actual duration, Test 1 mean estimated duration, and Test 2 mean estimated duration is striking in the case of many of the participants. Comparison of actual and Test 2 estimated durations indicated no significant difference [$F(1,143) = .0008, n.s.$], as did the comparison of the Test 1 and Test 2 estimates [$F(1,143) = .144, n.s.$]. Pearson correlations indicated significant relationships between actual duration and Test 2 estimates ($r = .61, p < .01$), and between Test 1 and Test 2 estimates ($r = .51, p < .01$).

Individual estimates were examined by a comparison of the Test 1 and Test 2 estimation errors. A signed error

Table 3
Tests 1 and 2 Mean Duration Estimates and Absolute Errors

| N | Mean Actual Duration | Mean | Mean | Test 1-Test 2 Absolute Difference |
|-------------|-------------------------|------------------------|------------------------|---|
| | | Estimated at Test 1 | Estimated at Test 2 | |
| 15 | 16.7 | 19.6 | 19.9 | .3 |
| 7 | 10.7 | 11.4 | 8.8 | .6 |
| 9 | 9.8 | 6.6 | 9.4 | 2.7 |
| 4 | 11.0 | 9.2 | 11.2 | 2.0 |
| 49 | 47.2 | 39.3 | 36.3 | 2.9 |
| 5 | 7.4 | 7.2 | 28.2 | 21.0 |
| 16 | 43.0 | 49.6 | 80.1 | 30.5 |
| 25 | 20.6 | 13.7 | 13.2 | .6 |
| 5 | 4.6 | 6.8 | 5.8 | 1.0 |
| 7 | 40.8 | 49.1 | 31.8 | 17.2 |
| 5 | 11.6 | 11.4 | 11.0 | .4 |
| Overall 147 | 29.8 | 27.2 | 29.8 | 3.5 |

score (actual duration – estimated duration) was calculated for both Test 1 and Test 2 estimates, where a negative score indicated overestimation and a positive score underestimation of actual duration. Inspection of the Test 1 and Test 2 signed errors indicated that they were identical for 22.9% of the estimates. Of the remaining events, 30.5% of the signed errors were smaller at Test 2. Thus the majority of estimates were either identical at Test 2 to that given at Test 1 or more accurate than that given at Test 1. This conclusion is reflected in the overall mean Test 2 signed error (.15 days), which is smaller than the overall Test 1 signed error (2.5 days), although comparison of these means indicated no significant difference [$F(1,143) = .144$, n.s.].

The consistency between the Test 1 and Test 2 estimates suggests that event memory ratings ($r = .15$) and retention interval ($r = .15$) were unlikely to be related to Test 2 estimated duration, and neither of the correlations was significant. Furthermore, Test 2 estimation errors (signed) were not related to event memory ratings ($r = -.04$), retention interval ($r = -.07$), distinctiveness of the event ($r = -.09$), or importance of the event ($r = -.12$). Participants' ratings of confidence in their Test 2 estimated event duration did not correlate significantly with estimation signed error ($r = -.14$) either.

DISCUSSION

The participants' memory for the study events showed a degree of forgetting, with just under half the events rated as forgotten. However, the participants were reasonably good at identifying that an event had actually happened, even if they could not remember it, since they assigned only a very small number of the true events to the "never happened" category. Thus, while participants clearly had trouble recalling some events, they were able to differentiate events that had happened from the false events. There was also some degree of relationship between the Test 1 and Test 2 memory data, with the scores formed from the Test 1 objectively determined memory categories correlating with the Test 2 memory ratings. This finding is encouraging for autobiographical memory research, which often measures memory strength through the use of rating data.

Interestingly, 4.9% ($n = 28$) of the events that were not remembered at Test 1 were rated as remembered at Test 2. Are these instances of recovered memories? Perhaps events had occurred between Test 1 and Test 2 that reestablished these memories (e.g., meeting up with an old friend who relates a number of past events that you had forgotten). Obviously, they were all plausible memories, and indeed the events had actually happened. This reasonable level of re-remembering was paralleled by a reasonably high level of acceptance of the false events as actually having happened—recall that 21.5% of the 514 false events were rated as remembered. As noted, a detailed analysis of the data relating to the false events is reported in Burt et al. (2000).

The hypothesis that estimating the duration of everyday events involves reconstruction is considerably bolstered by the results obtained in this study. A remarkable level of consistency was found in the duration estimates obtained at Test 1 and Test 2. Furthermore, both these sets of estimates were similar to actual event duration. While duration estimation models such as Ornstein's (1969) storage size hypothesis would predict that the passage of 10 years would significantly alter memory for event duration (estimates should decrease because of the forgetting of event content), a reconstructive model predicts the results that were obtained. Furthermore, the Test 2 memory ratings for the duration events were not significantly related to estimated duration, and the Test 2 memory ratings were on average significantly smaller than those obtained at Test 1, yet no systematic reduction in estimated duration was observed.

A particularly surprising result was the tendency toward backward telescoping evident in the participants' dating of their events. This result is in contrast to the many studies that have found remote events to be generally dated too recently, or forward telescoped (e.g., Bradburn et al., 1987; Burt, 1992b; Ferguson & Martin, 1983; Kemp, 1988, 1994, 1996; Means & Loftus, 1991; Rubin & Baddeley, 1989; Wright, Gaskell, & O'Muircheartaigh, 1997). Several explanations might be offered for this finding.

Distance-based models propose a link between date estimation and the strength of the event's memory (e.g., Bradburn et al., 1987; Brown et al., 1985; Friedman & Kemp, 1998; Kemp & Burt, 1998). However, we found no evidence of a relationship between rated event memory and signed dating error. In contrast, the boundary models of dating bias proposed by Huttenlocher, Hedges, and Prohaska (1988); Huttenlocher, Hedges, and Bradburn (1990); and Rubin and Baddeley (1989), which fall into the location-based category, might offer a partial explanation if it is assumed that the participants set up a recent boundary to their dating that was clearly remembered, and this had the effect of holding dates in the time preceding it. Two recent boundaries seem possible. The *true recent boundary* is the end date of the diary records that the participant submitted at Test 1. This date (e.g., December 31, 1987) is the last date from which an event could have been sampled, and no event could have occurred more recently than this date. Recall that only seven events, or 1.2%, were dated more recently than the true recent boundary. While the participants may have remembered this true recent boundary, it probably was not particularly unique or significant since many of the diarists were continuing their diary keeping during and after the Test 1 period. Perhaps a more memorable recent boundary was the date (or period of time, since some participants were interviewed over several sittings) at which Test 1 itself took place. The extensive examination of each participant's memory that occurred at Test 1 was probably a unique event in his/her life and the type of experience that might be expected to form a

landmark event. As a landmark event, the participant would be expected to be able to recall when it occurred. Furthermore, our initial contact with the diarists for Test 2 undoubtedly prompted some recollection of Test 1 and perhaps even some extended discussion with a partner of when and what had happened at Test 1. It might be reasonable to assume that the Test 1 recent boundary would be more memorable than the remote boundary (the date of the first diary entry in a participant's diary set), and thus create a very salient period in time before which all dates should be (and were) assigned. Of course, a recent boundary does not necessarily create backward telescoping, but it could reduce the likelihood of forward telescoping.

If the recent boundary is operating as we suggest, it might have worked in conjunction with other strategies that participants may use to date the events. An associative theory of dating was recently proposed by Kemp (1999), and this places considerable emphasis on the operation of contextual factors in date generation. Kemp's theory suggests that life events (memories) are linked in associative networks by contextual themes (e.g., living in a particular place or activities with a select group of people). Conway's (1992, 1995) work on autobiographical memory organization also emphasizes thematic structures. If an individual is asked to date an event (and cannot recall the actual date), part of his/her construction of the event's date may involve searching memory for events with a similar context for which some date information can be recalled.

To understand the nature of the search process for contextually associated events in our sample, it is necessary to consider the age of the participants when the events they were attempting to date occurred and their age at Test 2. The events they were dating occurred between the ages of 10 and 16 years. During this time the participants were probably living at home. Furthermore, the dated events may have more common characteristics with the events that occurred in the preceding years than with those that occurred in the following years. After 16 years of age the participants might have left home, moved to tertiary education, begun an active sex life, and so on. In contrast, the mean age of the participants at Test 2 was 30.8 years. It seems reasonable to conclude that the participants underwent major changes in the contextual associates of their life events between Test 1 and Test 2. At the least they have gone from adolescence to adulthood.

If the context of the dated events is unique because of the time in the participant's life when they occurred, any attempt to construct date information using contextual associates might occasion a search that takes the participant further away from the present than might be the case if, for example, he/she was dating events that had occurred in their late 20s. The slight increase in the use of landmark events between Test 2 and Test 1 that were seen as occurring before the events to be dated is consistent with this process. Thus backward telescoping may have occurred, at least in part, because the participants searched

memory for contextually similar events, for which they could access some date information, and this search tended to focus on an early period of life. Therefore the nature of dating error might be partially dependent on when in the life span an event occurred and when in the life span an individual is attempting to date the event. If, as was perhaps the case in this study, the two points in time were separated by a major life transition, then backward telescoping might occur. If, on the other hand, the time of event occurrence and the time of dating both occur in the same life period, forward telescoping might be expected.

The issue of forward and backward telescoping not only has interested event dating researchers, but has been discussed in relation to errors in survey work that focuses on activity frequency estimates (e.g., Bradburn et al., 1987; Neter & Waksberg, 1964; Penick & Owens, 1976; Schnieder & Sumi, 1981; Sudman & Bradburn, 1973, 1974, 1982). Forward telescoping has been linked to the tendency to overestimate activity frequency estimation because events that occurred outside the questions reference period are recalled as occurring within it (e.g., Bradburn et al., 1987). Backward telescoping has been noted when the reference period defined in the question is not bound at the recent end by the interview (e.g., *During 1987 how many times did you visit a doctor?*). In this type of a backward-telescoping error, events move *back* into the question's reference period. The results of the present study suggest that backward telescoping might see events moving out of a reference period (beyond the back boundary), creating underestimation of event frequency. Furthermore, this form of backward telescoping could be a particular problem if a survey were to question an adult population about activities undertaken in its teens.

In summary, the present study reports one of the few within-subjects repeated measures analyses of autobiographical memory. Duration estimation of everyday events seems very stable over time, and one might perhaps expect that this aspect of memory will remain constant over the entire life span. Dating errors, perhaps for the first time, have been shown to differ from the expected pattern. It appears that the conditions under which most studies of dating have operated have in part contributed to the error patterns that have been observed. The backward telescoping observed in this study seems reasonably well accounted for within the framework of the establishment of a recent boundary and Kemp's (1999) associative model of dating.

REFERENCES

- AVANT, L. L., & LYMAN, P. J. (1975). Stimulus familiarity modifies perceived duration in pre-recognition visual processing. *Journal of Experimental Psychology: Human Perception & Performance*, *1*, 205-213.
- BADDELEY, A. D., LEWIS, V., & NIMMO-SMITH, I. (1978). When did you last. . . ? In M. M. Gruneberg & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 77-83). London: Academic Press.
- BARCLAY, C. R., & WELLMAN, H. M. (1986). Accuracies and inaccuracies in autobiographical memories. *Journal of Memory & Language*, *25*, 93-103.

- BLOCK, R. A. (1978). Remembered duration: Effects of event and sequence complexity. *Memory & Cognition*, **6**, 320-326.
- BLOCK, R. A., & REED, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 656-665.
- BRADBURN, N. M., RIPS, L. J., & SHEVELL, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, **236**, 157-161.
- BROWN, N. R., RIPS, L. J., & SHEVELL, S. K. (1985). The subjective dates of natural events in very-long-term memory. *Cognitive Psychology*, **17**, 139-177.
- BROWN, N. R., SHEVELL, S. K., & RIPS, L. J. (1986). Public memories and their personal context. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 137-158). Cambridge: Cambridge University Press.
- BRUCE, D., & VAN PELT, M. (1989). Memories of a bicycle tour. *Applied Cognitive Psychology*, **3**, 137-156.
- BUFFARDI, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception & Psychophysics*, **10**, 292-294.
- BURNSIDE, W. (1971). Judgment of short time intervals while performing mathematical tasks. *Perception & Psychophysics*, **9**, 404-406.
- BURT, C. D. B. (1992a). Reconstruction of the duration of autobiographical events. *Memory & Cognition*, **20**, 124-132.
- BURT, C. D. B. (1992b). Retrieval characteristics of autobiographical memories: Event and date information. *Applied Cognitive Psychology*, **6**, 389-404.
- BURT, C. D. B. (1993). The effect of actual event duration and event memory on the reconstruction of duration information. *Applied Cognitive Psychology*, **7**, 63-73.
- BURT, C. D. B. (1999). The effect of categorization of action speed on estimated duration. *Memory*, **7**, 345-355.
- BURT, C. D. B., & KEMP, S. (1991). Retrospective duration estimation of public events. *Memory & Cognition*, **19**, 252-262.
- BURT, C. D. B., & KEMP, S. (1994). Construction of activity duration and time management potential. *Applied Cognitive Psychology*, **8**, 155-168.
- BURT, C. D. B., KEMP, S., & CONWAY, M. A. (2000). *Event attributes that define false event acceptance*. Manuscript in preparation.
- BURT, C. D. B., & POPPLE, J. S. (1996). Effects of implied action speed on estimation of event duration. *Applied Cognitive Psychology*, **10**, 53-63.
- BURT, C. D. B., WATT, S. C., MITCHELL, D. A., & CONWAY, M. A. (1998). Retrieving the sequence of autobiographical event components. *Applied Cognitive Psychology*, **12**, 321-338.
- CONWAY, M. A. (1992). A structural model of autobiographical memory. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 167-194). Dordrecht, The Netherlands: Elsevier.
- CONWAY, M. A. (1995). Autobiographical knowledge and autobiographical memories. In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67-93). Cambridge: Cambridge University Press.
- CROVITZ, H. F., & SCHIFFMAN, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, **4**, 517-518.
- DOUGLAS, J. W. B., & BLOMFIELD, J. M. (1956). The reliability of longitudinal surveys. *Millbank Memorial Fund Quarterly*, **34**, 227-252.
- EKMAN, G., FRANKENHÄUSER, M., LEVANDER, S., & MELLIS, I. (1966). The influence of intensity and duration of electrical stimulation on subjective variables. *Scandinavian Journal of Psychology*, **7**, 58-64.
- FERGUSON, R. P., & MARTIN, P. (1983). Long-term temporal estimation in humans. *Perception & Psychophysics*, **33**, 585-592.
- FITZGERALD, J. M. (1981). Autobiographical memory reports in adolescence. *Canadian Journal of Psychology*, **35**, 69-75.
- FITZGERALD, J. M., & LAWRENCE, R. (1984). Autobiographical memory across the life-span. *Journal of Gerontology*, **39**, 692-699.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, **35**, 1-36.
- FRIEDMAN, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, **113**, 44-66.
- FRIEDMAN, W. J. (1996). Distance and location processes in memory for the times of past events. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 35, pp. 1-41). San Diego: Academic Press.
- FRIEDMAN, W. J., & KEMP, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development*, **13**, 335-367.
- FRIEDMAN, W. J., & WILKINS, A. J. (1985). Scale effects in memory for the time of events. *Memory & Cognition*, **13**, 168-175.
- HOLDING, D. H., NOONAN, T. K., PFAU, H. D., & HOLDING, C. S. (1986). Date attribution, age and the distribution of life-time memories. *Journal of Gerontology*, **41**, 481-485.
- HUTTENLOCHER, J., HEDGES, L., & BRADBURN, N. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 196-213.
- HUTTENLOCHER, J., HEDGES, L., & PROHASKA, V. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review*, **95**, 471-484.
- KEMP, S. (1987). Gestalt grouping effects in locating past events on time-lines. *Acta Psychologica*, **64**, 139-149.
- KEMP, S. (1988). Dating recent and historical events. *Applied Cognitive Psychology*, **2**, 181-188.
- KEMP, S. (1994). Bias in dating news and historical events. *Acta Psychologica*, **86**, 69-87.
- KEMP, S. (1996). Association as a cause of dating bias. *Memory*, **4**, 131-143.
- KEMP, S. (1999). An associative theory of estimating past dates and past prices. *Psychonomic Bulletin & Review*, **6**, 41-56.
- KEMP, S., & BURT, C. D. B. (1998). The force of events: Cross-modality matching the recency of news events. *Memory*, **6**, 297-306.
- KOWAL, K. H. (1987). Apparent duration and numerosity as a function of melodic familiarity. *Perception & Psychophysics*, **42**, 122-131.
- LARSEN, S. F., & THOMPSON, C. P. (1995). Reconstructive memory in the dating of personal and public news events. *Memory & Cognition*, **23**, 780-790.
- LINTON, M. (1975). Memory for real world events. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 376-404). San Francisco: Freeman.
- LINTON, M. (1978). Real world memory after six years: An in vivo study of very long term memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 69-76). London: Academic Press.
- LOFTUS, E. F., & FATHI, D. C. (1985). Retrieving multiple autobiographical memories. *Social Cognition*, **3**, 280-295.
- MCCONCHIE, R. D., & RUTSCHMANN, J. (1971). Human time estimation: On differences between methods. *Perceptual & Motor Skills*, **32**, 319-336.
- MCCORMACK, P. (1979). Autobiographical memory in the aged. *Canadian Journal of Psychology*, **33**, 118-124.
- MEANS, B., & LOFTUS, E. F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology*, **5**, 297-318.
- MEDNICK, S. A., & SHAFFER, J. B. P. (1963). Mothers' retrospective reports in child-rearing research. *American Journal of Orthopsychiatry*, **33**, 457-461.
- NETER, J., & WAKSBERG, J. (1964). A study of response errors in expenditure data from household interviews. *Journal of the American Statistical Association*, **59**, 18-55.
- ORNSTEIN, R. E. (1969). *On the experience of time*. Harmondsworth: Penguin.
- PENICK, B. K. E., & OWENS, M. E. B. (Eds.) (1976). *Surveying crime*. Washington, DC: National Academy of Sciences.
- PERLMUTTER, M., METZGER, R., MILLER, K., & NEZWORSKI, T. (1980). Memory of historical events. *Experimental Aging Research*, **6**, 47-60.
- PILLEMER, D. B., GOLDSMITH, L. R., PANTER, A. T., & WHITE, S. H. (1988). Very long-term memories of the first year in college. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 709-715.
- PILLEMER, D. B., RHINEHART, E. D., & WHITE, S. H. (1986). Memories of life transitions: The first year in college. *Human Learning: Journal of Practical Research & Applications*, **5**, 109-123.

- POYNTER, W. D., & HOMA, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, **33**, 548-560.
- PYLES, M. K., STOLZ, H. R., & MACFARLANE, J. W. (1935). The accuracy of mothers' reports on birth and developmental data. *Child Development*, **6**, 165-176.
- RACHLIN, H. C. (1966). Scaling subjective velocity distance and duration. *Perception & Psychophysics*, **1**, 77-82.
- ROBINSON, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology*, **8**, 578-595.
- ROBINSON, J. A. (1986). Autobiographical memory: A historical prologue. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 159-188). Cambridge: Cambridge University Press.
- RUBIN, D. C., & BADDELEY, A. D. (1989). Telescoping is not time compression: A model of the dating of autobiographical events. *Memory & Cognition*, **17**, 653-661.
- SCHIFFMAN, H. R., & BOBKO, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology*, **103**, 156-159.
- SCHIFFMAN, H. R., & BOBKO, D. J. (1977). The role of number and familiarity of stimuli in the perception of brief temporal intervals. *American Journal of Psychology*, **90**, 85-93.
- SCHNEIDER, A. L., GRIFFITH, W. R., SUMI, D. H., & BURCART, J. M. (1978). *Portland forward records check of crime victims*. Washington, DC: U. S. Department of Justice.
- SCHNIEDER, A. L., & SUMI, D. (1981). Patterns of forgetting and telescoping: An analysis of LEAA survey victimization data. *Criminology*, **19**, 400-410.
- SHUM, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin*, **124**, 423-442.
- SUDMAN, S., & BRADBURN, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, **68**, 805-815.
- SUDMAN, S., & BRADBURN, N. M. (1974). *Response effects in surveys*. Chicago: Aldine.
- SUDMAN, S., & BRADBURN, N. M. (1982). *Asking questions*. San Francisco: Jossey-Bass.
- THOMAS, E. A., & BROWN, I. (1974). Time estimated and the filled-duration illusion. *Perception & Psychophysics*, **16**, 449-458.
- THOMPSON, C. P. (1982). Memory for unique personal events: The roommate study. *Memory & Cognition*, **10**, 324-332.
- THOMPSON, C. P. (1985a). Memory for unique personal events: Effects of pleasantness. *Motivation & Emotion*, **9**, 277-289.
- THOMPSON, C. P. (1985b). Memory for unique personal events: Some implications of the self-schema. *Human Learning: Journal of Practical Research & Applications*, **4**, 267-280.
- THOMPSON, C. P., SKOWRONSKI, J. J., & BETZ, A. L. (1993). The use of partial temporal information in dating personal events. *Memory & Cognition*, **21**, 352-360.
- THOMPSON, C. P., SKOWRONSKI, J. J., LARSEN, S. F., & BETZ, A. L. (1996). *Autobiographical memory: Remembering what and remembering when*. Mahwah, NJ: Erlbaum.
- THOMPSON, C. P., SHOWRONSKI, J. J., & LEE, D. J. (1988). Telescoping in dating naturally occurring events. *Memory & Cognition*, **16**, 461-468.
- WAGENAAR, W. A. (1986). My memory: A study of autobiographical memory over 6 years. *Cognitive Psychology*, **18**, 225-252.
- WHITE, R. T. (1982). Memory for personal events. *Human Learning*, **1**, 177-183.
- WRIGHT, D. B. (1998). Modeling clustered data in autobiographical memory research: The multilevel approach. *Applied Cognitive Psychology*, **12**, 339-357.
- WRIGHT, D. B., GASKELL, G. D., & O'MUIRCHARTAIGH, C. A. (1997). Temporal estimation of major news events: Re-examining the accessibility principle. *Applied Cognitive Psychology*, **11**, 35-46.

(Manuscript received September 9, 1999;
revision accepted for publication April 13, 2000.)